

Art Unit: 2162

### **DETAILED ACTION**

1. Applicant has amended claims 1, 11, 27, 34, 43 in the amendment filed on 3/28/2008.

Claims 1-38, 41-44 are pending in this Office Action.

### **EXAMINER'S AMENDMENT**

2. An examiner's amendment to the record appears below. Should the changes and/or additions be unacceptable to applicant, an amendment may be filed as provided by 37 CFR 1.312. To ensure consideration of such an amendment, it MUST be submitted no later than the payment of the issue fee.

Authorization for this examiner's amendment was given in a telephone interview with Attorney John W. Branch on 4/22/2008.

#### **In claims:**

Please replace claims 1, 4, 8, 9, 11-14, 17, 20, 25, 27, 34 and 37 with amended 1, 4, 8, 9, 11-14, 17, 20, 25, 27, 34 and 37 .

Please cancel claims 5-7, 15-16, 18-19, 21 and 42-43.

Art Unit: 2162

1. (Currently Amended) A method for extracting data from a network by a server, comprising:

a) creating a database-structured query with at least one fundamental clause including a web domain address used for locating the data, based, in part, on a user input, wherein the database-structured query includes a regular expression used to determine the data to extract, and a conditional expression describing a number where to start searching and another number where to stop searching the data at the web domain address, wherein the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user;

b) creating a template of the regular expression used to extract the data;

c) providing authentication data to the web domain address;

d) determining the web domain address on the network from which to extract the data;

e) extracting the data from the web domain address directly by retrieving a non-database structured arrangement of data from the determined web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting data from the web domain further comprises matching a plurality of patterns contained within the regular expression to retrieved data to determine the data to extract,

(f) repeating steps (d) and (e) in an iterative manner based on the at least one fundamental clause;

(g) reshaping the extracted data to a predetermined format ; and

Art Unit: 2162

(h) providing the extracted data from the determined web domain address, wherein the extracted data is provided in a tab delimited data file, and wherein the tab delimited data file is provided directly to the user.

4.(Currently amended) The method of Claim 2, wherein directly extracting data from the determined web domain address, further comprises, following links contained within the web domain address until the links have been exhausted or following the links until a predetermined limit is reached.

8. (Currently amended) The method of Claim 1, wherein directly extracting the data from the web domain address, further comprises: reducing the retrieved data to a region of interest; and searching the region of interest for the data matching a predetermined regular expression.

9.(Currently amended) The method of Claim 8, wherein directly extracting the data from the web domain address, further comprises, storing data matching the regular expression.

Art Unit: 2162

11. (Currently Amended) A computer-readable storage device having computer-executable instructions for extracting data from a network, the computer-executable instructions enabling actions comprising:

a) creating a database-structured query with at least one fundamental clause including a web domain address used for locating the data, based, in part, on a user input, wherein the database-structured query includes a regular expression used to determine the data to extract, and a conditional expression describing a number where to start searching and another number where to stop searching the data at the web domain address, wherein the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user;

b) creating a template of the regular expression used to extract the data;

c) providing authentication data to the web domain address;

d) determining the web domain address on the network from which to extract the data;

e) extracting the data from the web domain address directly by retrieving a non-database structured arrangement of data from the determined web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting data from the web domain further comprises matching a plurality of patterns contained within the regular expression to retrieved data to determine the data to extract.

(f) repeating steps (d) and (e) in an iterative manner based on the at least one fundamental clause;

Art Unit: 2162

(g) reshaping the extracted data to a predetermined format ; and

(h) providing the extracted data from the determined web domain address, wherein the extracted data is provided in a tab delimited data file, and wherein the tab delimited data file is provided directly to the user.

12. (Currently Amended) The computer-readable storage device of Claim 11, wherein the database-structured query, further comprises, a network address included within the database-structured query indicating a starting point to start to crawl the network.

13. (Currently Amended) The computer-readable storage device of Claim 12, wherein the network address, further comprises at least one universal resource locator (URL) string generated by a sequence or list function.

14.(Currently Amended) The computer-readable storage device of Claim 11, wherein the web domain address, further comprises at least one link to another web domain address for locating data to extract.

17. (Currently Amended) A system for extracting data from a network comprising:

a client computer system having a hardware computing device and a client network connection to the network and communicating with a server computer system, the client computer system creating a database-structured

Art Unit: 2162

query with at least one fundamental clause, based, in part, on a user input, wherein the database-structured query includes a regular expression used to determine the data to extract and a conditional expression describing a number where to start searching and another number where to stop searching the data at a web domain address, wherein the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user;

an editor for creating a template of the regular expression used to extract the data;

the server computer system having a server network connection to the network and communicating with the client computer system, the server computer system further performs actions, comprising:

receiving the database-structured query from the client computer system;

determining the web domain address on the network from which to extract at least a portion of the data relevant to the query, wherein the determined web domain address is provided by the database-structured query;

providing authentication data to the web domain address;

extracting directly at least the portion of the data from the web domain address by retrieving a non-database structured arrangement of data from the determined web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting the portion of data from the web domain further comprises matching a

Art Unit: 2162

plurality of patterns contained within the regular expression to retrieved data to determine the data to extract;

repeating, based on the at least one fundamental clause, actions of the determining a web domain address and the extracting the portion of data from the web domain address in an iterative manner;

reshaping the extracted data to a predetermined format; and

providing the extracted data from the web domain address, wherein the extracted data is provided in a tab delimited data file, and wherein the tab delimited data file is provided directly to the user.

20. (Currently amended) The system of Claim 17, wherein the regular expression within the database-structured query, further comprises, at least one symbol used to denote a part of the regular expression that a user desires to extract.

25. (Currently amended) The computer-readable storage device of claim 11, wherein the web domain address further comprises at least one link address that is followed to locate data to extract until a user-specified number of links is reached.

27. (Currently Amended) A method of extracting data from a network by a server, comprising:

(a) creating a database-structured query with at least one fundamental clause including a web domain address at the server based, in part, on a user

Art Unit: 2162

input, wherein the database-structured query includes a regular expression used to determine the data to extract, creating a template of the regular expression used to extract the data, providing authentication data to the web domain address;

(b) determining a website to search based in part on the database-structured query and wherein the database-structured query further includes a conditional expression describing a number where to start searching and another number where to stop searching the data at the website, wherein the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user;

(c) extracting the data at the website directly by retrieving a non-database structured arrangement of data from the web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting data from the web domain address further comprises matching a plurality of patterns contained within the regular expression to retrieved data to determine the data to extract, wherein the website is processed as a searchable database;

(d) repeating steps (b) and (c) in an iterative manner based on the at least one fundamental clause; and

(e) reshaping the extracted data to a predetermined format;

(f) providing the extracted data from the website, wherein the extracted data is provided in a data log, and wherein the data log is provided directly to the user.



34. (Currently Amended) A method of extracting data within at least one webpage, comprising:

(a) generating a database-structured query with at least one fundamental clause including a web domain address based, in part, on a user's input, wherein the database-structured query further includes a regular expression used to determine the data to extract and a conditional expression describing a number where to start searching and another number where to stop searching the data at the at least one webpage, wherein the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user;

(b) creating a template of the regular expression used to extract the data and providing authentication data to the web domain address;

(c) determining the at least one webpage with the data, wherein the determination of the at least one webpage is provided by the database-structured query;

d) ) extracting the data at the at least one webpage directly by retrieving a non-database structured arrangement of data from the web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting data from the web domain further comprises matching a plurality of patterns contained within the regular expression to retrieved data to determine the data to extract, wherein the extracted data that satisfies a query condition includes at least one binary file;

Art Unit: 2162

(e) repeating steps (c) and (d) in an iterative manner based on the at least one fundamental clause;

(f) reshaping the extracted data to a predetermined format; and

(g) providing the extracted data from the at least one webpage, wherein the extracted data is provided in a data log, and wherein the data log is provided directly to the user.

37. (Currently amended) The method of Claim 34, wherein the structured query is generated to parse a limited portion of the data of the at least one webpage with limits predetermined by the user.

***Allowable Subject Matter***

3. Claims 1-4, 8-14, 17, 20, 22-38, 41, 44 are allowed.

The prior art of record alone or in combination, does not teach or fairly suggest the combination of steps as recited in independent claims 1, 11 and 17 (system claim), wherein “extracting the data from the web domain address directly by retrieving a non-database structured arrangement of data from the determined web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting data from the web domain further comprises matching a plurality of patterns contained within the regular expression to the retrieved data to determine the data to extract, wherein the database-structured query includes a regular expression used to determine the data to extract, and a conditional expression describing a number where to start searching and another number to stop searching the data at the web domain address, the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user; providing authentication data to the web domain address”;

The prior art of record alone or in combination, does not teach or fairly suggest the combination of steps as recited in independent claim 27, wherein “extracting the data at the webpage directly by retrieving a non-database structured arrangement of data from the web domain address and performing the database-structured query upon the retrieved non-database structured

Art Unit: 2162

arrangement of data, wherein the extracting data from the web domain further comprises matching a plurality of patterns contained within the regular expression to retrieved data to determine the data to extract, wherein the extracted data that satisfies a query condition includes at least one binary file, wherein the database-structured query further includes a regular expression used to determine the data to extract and a conditional expression describing a number where to start searching and another number where to stop searching the data at the at least one webpage, the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user; creating a template of the regular expression used to extract the data and providing authentication data to the web domain address”; and

The prior art of record alone or in combination, does not teach or fairly suggest the combination of steps as recited in independent claim 34, wherein “extracting the data at the webpage directly by retrieving a non-database structured arrangement of data from the web domain address and performing the database-structured query upon the retrieved non-database structured arrangement of data, wherein the extracting data from the web domain further comprises matching a plurality of patterns contained within the regular expression to retrieved data to determine the data to extract, wherein the extracted data that satisfies a query condition includes at least one binary file, wherein the database-structured query further includes a regular expression used to determine the data to extract and a conditional expression describing a number where to start searching and another number where to stop searching

Art Unit: 2162

the data at the at least one webpage, the number where to start searching, the another number where to stop searching and an amount to increment are defined by a user; creating a template of the regular expression used to extract the data and providing authentication data to the web domain address”.

The dependent claims, bring definite, further limiting, and fully enabled by the specification are also allowed.

### **Contact Information**

4. Any inquiry concerning this communication or earlier communications from the examiner should be directed to Cam Y T. Truong whose telephone number is (571) 272-4042. The examiner can normally be reached on Monday to Friday.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, John Breene can be reached on (571) 272-4107. The fax phone number for the organization where this application or proceeding is assigned is 571-273-8300.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free). If you would like assistance from a USPTO Customer Service Representative or access to the automated information system, call 800-786-9199 (IN USA OR CANADA) or 571-272-1000.

/Cam Y Truong/  
Primary Examiner, Art Unit 2162